TRUNCATED COUNTING FUNCTIONS OF HOLOMORPHIC CURVES IN ABELIAN VARIETIES

Ryoichi Kobayashi

Graduate School of Mathematics, Nagoya University

Abstract. A new proof of the Second Main Theorem with truncation level 1 for Zariski-dense holomorphic curves into Abelian varieties, which has just been proved by Yamanoi [Y2], is presented. Our proof is based on the idea of the "Radon transform" introduced in [K2] combined with consideration on certain singular perturbation of the probability measures on the parameter space which appears in the "Radon transform".

1. Introduction.

The Nevanlinna Theory describes "intersection theory" of holomorphic curves from $\mathbb{C}$ and divisors in a projective algebraic variety in terms of the relationship among "basic functions". Let us recall the definition of the basic functions in Nevanlinna theory. Let X be a smooth projective variety, D an effective divisor, σ a canonical section of $\mathcal{O}_X(D)$ and $\|\cdot\|$ a smooth Hermitian norm of $\mathcal{O}_X(D)$. Let $f : \mathbb{C} \to X$ be a holomorphic curve such that $f(\mathbb{C}) \not\subset \mathrm{Supp}(D)$. The "intersection theory" described by the Nevanlinna theory of (transcendental) holomorphic curves and divisors has two aspects. One is to measure how a holomorphic curve can approximate a given divisor. The other is to measure how often a holomorphic curve can intersect a divisor. We measure the approximation of a holomorphic curve $f : \mathbb{C} \to X$ by the asymptotic behavior of the **proximity function**

$$m_{f,D}(r) = \int_0^{2\pi} \log^+ \frac{1}{\|\sigma \circ f(re^{i\theta})\|} \frac{d\theta}{2\pi} \ .$$

Note that $\|\sigma \circ f(re^{i\theta})\|$ is equivalent to the Euclidean distance between $f(re^{i\theta})$ and D. Let $n_{f,D}(t)$ (resp. $n_{f,D}(0)$) denote the numbers of zeros of $\sigma \circ f$ in $\mathbb{C}(t) = \{z \in \mathbb{C} ; |z| < t\}$ (resp. $\deg_0(\sigma \circ f)$). We then measure how often f intersects D by the asymptotic behavior of the **counting function**

$$\begin{aligned} N_{f,D}(r) &= \sum_{0 \neq a \in \mathbb{C}(r)} \deg_a(\sigma \circ f) \log \left| \frac{r}{a} \right| + n_{f,D}(0) \log r \\ &= \int_0^r \frac{n_{f,D}(t) - n_{f,D}(0)}{t} dt + n_{f,D}(0) \log r \ . \end{aligned}$$

The sum of these two functions is the **height function**

$$T_{f,D}(r) = m_{f,D}(r) + N_{f,D}(r) \ .$$

We call these three functions the “basic functions” in Nevanlinna theory.

The First Main Theorem in Nevanlinna Theory states that the asymptotic behavior of the height function, considered as an element of

$$\frac{\text{all functions on } \mathbb{R}_{>0}}{\text{all bounded functions on } \mathbb{R}_{>0}}$$

depends only on the linear equivalence class of the divisor D. Indeed, if $D_1 = (\sigma_1) \sim D_2 = (\sigma_2)$ (linearly equivalent), then the difference of $T_{f,D_1}(r) - T_{f,D_2}(r)$ essentially depends on the leading coefficient of the Laurent expansion of the meromorphic function $\frac{\sigma_1}{\sigma_2} \circ f$ at $z = 0$.

There is another kind of counting function, namely, the **ramification counting function**. Let R be a Riemann surface and $f : \mathbb{C} \to R$ a holomorphic map. Then the ramification counting function is defined as

$$N_{f,\mathrm{Ram}}(r) = \sum_{0 \neq a \in \mathbb{C}(r)} (\mathrm{mult}_a(f) - 1) \log \left| \frac{r}{a} \right| + (\mathrm{mult}_0(f) - 1) \log r$$

where $\mathrm{mult}_a(f)$ is the multiplicity of f at $z = a$ (i.e., 1 plus the vanishing order of the derivative f' at $z = a$). This coincides with the usual counting function of the solutions of the equation $f'(z) = 0$ counted with multiplicities. In this article, the ramification counting function plays an essential role.

The estimate of the Second Main Theorem type is most important in Nevanlinna Theory. For instance, the Conjecture 1.1 is a typical example of the leading conjecture in the Nevanlinna Theory (see, for instance, [L1,2], [NoO], [V1]):

Conjecture 1.1. *Let X be a smooth projective variety, D a divisor with at worst simple normal crossings and $E \to X$ any ample line bundle. Let ε be any positive number. Then there exists a proper algebraic subset $Z = Z(X, D, \varepsilon, E)$ with the following property. Let $f : \mathbb{C} \to X$ be any holomorphic curve such that $f(\mathbb{C}) \not\subset$* Supp$(D)$. *Then if f is algebraically non-degenerate in the sense that the the image of f is Zariski dense, we have*

$$m_{f,D}(r) + N_{W(f),0}(r) + T_{f,K_X}(r) \leq \varepsilon T_{f,E}(r) /\!/_{\varepsilon} \ ,$$

where $N_{W(f),0}(r)$ is the counting function for the equation $W(f) = 0$, where $W(f)$ is a conjectural object interpreted as a kind of “Wronskian” determinant of the holomorphic curve $f : \mathbb{C} \to X$. Moreover, if f is algebraically non-degenerate in the weaker sense that $f(\mathbb{C}) \not\subset Z$ holds, then there is a modification of $W(f)$ so that the inequality of the same form with modified $W(f)$ hold.

The term $\varepsilon T_{f,E}(r)$ is an error term and $/\!/_{\varepsilon}$ means that the inequality holds outside an ε-dependent Borel set of ε-dependent finite Lebesgue measure. In general, the counting function $N_{W(f),0}(r)$ should generalize the ramification counting function for holomorphic maps to Riemann surfaces.

The following list essentially exhausts all known cases where Conjecture 1.1 is true.

(i) [Nevanlinna, Ahlfors] The case X is a compact Riemann surface ([N], [A]). In this case, $Z = \emptyset$.
(ii) [Nevanlinna-Cartan theory and its generalization] The case $X = \mathbb{P}^n(\mathbb{C})$, D a collection of hyperplanes in general position and the holomorphic curve $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ is linearly non-degenerate. ([C], [W] , [A], [F] and [V2,3]). In [V2,3] Vojta proved that there exists a (effectively computable) finite collection of proper linear sybspaces Z such that if $f(C)$ is not contained in Z, the inequality of Conjecture 1.1 holds for linearly non-degenerate f and even for linearly degenerate f such that $f(\mathbb{C}) \not\subset Z$ the same inequality without the Wronskian term (i.e., the term like $N_{W(f),0}(r)$) holds. It would be an interesting problem to recover a "Wronskian term" in the linearly degenerate case. Recently, Ru [R] proved a defect relation when D is a general hypersurface configuration.
(iii) [Bloch, Ochiai] The case that X is a subvariety (of general type) of an Abelian variety ([O], [GG], [NoO], [KS], [Y2]). In this case, D is empty and Z is the union of all translations of proper Abelian subvarieties contained in X.
(iv) The case that X is an Abelian variety and D is an arbitrary divisor ([SY], [Y1,2], [NoWY], [M], [K2]). In this case $Z = \emptyset$.

These results are called the Second Main Theorem. The most well-understood Second Main Theorem for holomorphic curves into higher dimensional targets is the case (ii) where $X = \mathbb{P}^n(\mathbb{C})$ and D a collection of q linear divisors in general position. In this case $W(f)$ is the usual Wronskian determinant of f with respect to the canonical affine local coordinate system of $\mathbb{P}_n$. Replacing $\deg_a(\sigma \circ f)$ by $\min\{\deg_a(\sigma \circ f), k\}$ in the definition of the usual counting function, we get the counting function truncated at level k, which we denote by $N_{k,f,D}(r)$. Define the residual counting function $N^k_{f,D}(r)$ with truncation level k by

$$N^k_{f,D}(r) = N_{f,D}(r) - N_{k,f,D}(r)\ .$$

So $N^k_{f,D}(r)$ counts only intersections of f and D with multiplicity $m \geq k$ with weight $m-k$ (i.e., replacing $\deg_a(\sigma \circ f)$ by $\max\{\deg_a(\sigma \circ f) - k, 0\}$ in the definition of the usual counting function). We then have

$$N^l_{f,D}(r) \leq N^k_{f,D}(r)$$

if $k \leq l$. Restricting the equation $W(f) = 0$ to D, we have

$$N^n_{f,D}(r) = N_{W(f),0}(r)\ .$$

Therefore, if $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ is linearly non-degenerate, then the Second Main Theorem with the truncated counting function of truncation level n holds:

$$m_{f,D}(r) - (n+1)T_{f,\mathcal{O}(1)}(r) + N^n_{f,D}(r) \leq O(\log^+(rT_{f,H}(r)))/\!/\ .$$

It follows from the above argument that the reason for the truncation level being $n = \dim \mathbb{P}^n(\mathbb{C})$ comes from the key role played by the Wronskian determinant $W(f)$. In general case, we conjecture that the Wronskian $W(f)$ must be replaced by its generalization (still some kind of Wronskian), and if so, the expected truncation level will be $n = \dim X$.

On the other hand, we expect some special property on the truncated counting function, if the target X is a special variety, for instance, if X is an Abelian variety. In fact, the theory of theta divisors of Jacobian varieties of compact Riemann surfaces suggests that the value distribution of holomorphic curves into the Jacobian variety of a compact Riemann surface should have some resemblance to the value distribution theory with one-dimensional targets. For instance, it is natural to ask if the relevant truncation level is one for holomorphic curves in Jacobian varieties (more generally, holomorphic curves in Abelian varieties).

Inspired by works of McQuillan [M] and Brunella [B], Yamanoi [Y2] discovered that the truncation level is taken to be one in the Second Main Theorem for *Zariski-dense* holomorphic curves $f : \mathbb{C} \to A$ into an Abelian variety A (D being any divisor):

$$m_{f,D}(r) + N^{1}_{f,D}(r) \leq \varepsilon T_{f,E}(r) /\!/_{\varepsilon} \ . \tag{1}$$

The assumption f being Zariski-dense cannot be removed. Moreover, the error term is of the form $\varepsilon T_{f,E}(r)$ (ε being any positive number) and this is not improved to the form $O(\log^{+}(rT_{f,E}(r)))$ ([Y2]). As an application of this result, Yamanoi [Y2] gave a new proof of the Bloch-Ochiai Theorem [O].

The purpose of this article is to give a new proof of Yamanoi's result (1) ([Y2]) by using the method of "Radon transform" developed in [K2]. In the course of the proof, we will clarify a simple reason[1] why the truncation level in the Second Main Theorem is taken to be one for Zariski-dense holomorphic curves into Abelian varieties.

Our proof is separated into two steps.

In the first step, we introduce the idea of the "Radon transform" of holomorphic curves in Abelian varieties ([K2]) in order to study the intersection of a given holomorphic curve $f : \mathbb{C} \to A$ with an ample divisor D. The "Radon transform" transforms a given entire holomorphic curve $f : \mathbb{C} \to A$ into a family of holomorphic maps $\{f_\lambda : Y_\lambda \to S_\lambda\}_{\lambda \in \Lambda}$, where Y_λ is a finite analytic covering with projection $\pi_\lambda : Y_\lambda \to \mathbb{C}$ and $\{S_\lambda\}_{\lambda \in \Lambda}$ form a family of algebraically equivalent algebraic curves in A all passing through the neutral point. where Λ is the parameter space. To get information independent of λ, we apply the Second Main Theorem for holomorphic curves from finite covering of $\mathbb{C}$ to Riemann surfaces (see, for instance, [No]) to each $f_\lambda : Y_\lambda \to S_\lambda$ and "average" them over the parameter space Λ with respect to some probability measure[2]. In §2, we introduce the "Gaussian divisor" $D^{[1]}_\lambda \subset \mathbb{P}(TA)$ (resp. the "incidence divisor" $D_\lambda \subset A$) which are associated to the pair (D, S_λ), which is related to the ramification of $f_\lambda : Y_\lambda \to S_\lambda$ (resp. $\pi_\lambda : Y_\lambda \to \mathbb{C}$). Then D, D_λ, $D^{[1]}_\lambda$ and $H^{[1]}$ (the relative hyperplane bundle of $\mathbb{P}(TA)$) fits into the generalized Hurwitz formula (see (2) in §2). We introduce a notion "variation" of probability spaces $(\Lambda, d\lambda_t)$ $(0 < t < r)$ with certain "measure concentration", over which Nevanlinna theoretic functions containing λ as a parameter are "averaged".

[1] Definition 2.1-2.2 of the Radon transform and its property (response under certain singular perturbation of the probability space Λ involved in the Radon transform), which is presented in Lemmas 2.5 and 2.7.

[2] Which probability measure we should choose is essential in this article.

By examining the response of various counting functions under the "measure concentration", we know the special feature of "averaging over $(\Lambda, d\lambda_t)$" procedure. For any $\varepsilon > 0$, there exists a variation of probability spaces $(\Lambda, d\lambda_t)$ such that:

- The "average" of $N_{f_\lambda,\mathrm{Ram}}(r) - N_{f^{[1]},D_\lambda^{[1]}}(r)$ is negligible.

- The "average" of $N_{\pi_\lambda,\mathrm{Ram}}(r) - N_{f,D_\lambda}(r)$ is equal to $-N^1_{f,D}(r)$ modulo the small error of magnitude $\varepsilon\, T_{f,D}(r)$.

Combining these estimates with the generalized Hurwitz formula give rise to a Diophantine inequality (1) of the Second Main Theorem type, which is much stronger than the version with generally expected counting function $N^n_{f,D}(r)$ in the Second Main Theorem for holomorphic curves into n-dimensional Abelian varieties.

The second step, where we still use special properties of Abelian targets, consists of showing that the proximity function of the jet lifts of a Zariski-dense holomorphic curve $f : \mathbb{C} \to A$ with respect to the jet spaces of D is small in the sense that it is dominated by the quantity $\varepsilon\, T_{f,E}(r)$ (for any $\varepsilon > 0$ and any ample line bundle $E \to A$). Nevanlinna's lemma on logarithmic derivative (see Lemma 3.1) applied to the higher jet lifts of $f : \mathbb{C} \to A$ is essential in the argument of the second part.

The essential steps in our proof of (1) are Lemmas 2.5 and 2.7. To generalize these Lemmas to more general "Radon transform" defined for arbitrary holomorphic curves into general projective algebraic varieties seems to be promissing because of the rich flexibility in the definition of the "Radon Transform" (see [K3]). We therefore pose the following conjecture for future study:

Conjecture 1.2. *Let X, D, E and ε be as in Conjecture 1.1. Let $f : \mathbb{C} \to X$ be any holomorphic curve which is algebraically non-degenerate in the sense that the the image of f is Zariski dense in X. Then, we have*

$$m_{f,D}(r) + N^1_{f,D}(r) + N_{W(f),0}(r) + T_{f,K_X}(r) \leq \varepsilon\, T_{f,E}(r)/\!\!/_\varepsilon\ ,$$

where $N_{W(f),0}(r)$ is a "conjectural" counting function for the equation $W(f) = 0$, where $W(f)$ is a "conjectural" object interpreted as a kind of "Wronskian" determinant of the holomorphic curve $f : \mathbb{C} \to X$. Moreover, there exists a proper algebraic subset $Z = Z(X, D, \varepsilon, E)$ with the following property. If $f : \mathbb{C} \to X$ is any holomorphic curve satisfying the property $f(\mathbb{C}) \not\subset Z$, then there is a modification of $W(f)$ so that the same inequality as above should hold.

2. Radon Transformation.

Let A be an Abelian variety and D a reduced divisor. Suppose that D is not ample. Then there exists an Abelian variety A' of lower dimension, an ample divisor $D' \subset A'$ and a surjective morphism $\pi : A \to A'$ such that $\pi^* D' = D$. Therefore, the intersection theory of $f : \mathbb{C} \to A$ and D is reduced to that of $\pi \circ f : \mathbb{C} \to A'$ with D'. By this reason, we put the following assumptions:

- *Assumption 1.* The divisor D is ample and reduced.

Moreover, we assume the following

- The divisor D has at worst normal crossings.

This assumption is not necessary for the proof of (1) in the sense that our arguments can be modified to cover the cases without this. However, this assumption makes the arguments in this article considerably simple.

Let $\{S_\lambda\}_{\lambda \in \Lambda}$ (Λ being the parameter space) be an algebraic family of algebraically equivalent curves in A all passing through the neutral point p of A.

- *Assumption 2.* There is a Zariski open subset Λ^0 of Λ such that for all $\lambda \in \Lambda^0$ the curve S_λ is non-singular.

- *Assumption 3.* There is a positive integer k with the following property. The natural rational map $t : \Lambda \cdots \to A_p^{[k]}$ from Λ to the k-th projective jet space $A_p^{[k]}$ at p is surjective and generically finite.

The k-th projective jet space $A_p^{[k]}$ at the neutral point $p \in A$ is defined as follows: we start with the k-times iterated projective tangent bundle of A, i.e., the space obtained from A by iterating the operation of taking projective tangent bundle k-times. Then, $A_p^{[k]}$ consists of projective classes of k-jets of all germs of holomorphic curves passing through the neutral point p. The natural map t is defined by

$$\Lambda \ni \lambda \mapsto t(\lambda) := (S_\lambda^{[k]})_p \in A_p^{[k]} \ .$$

We introduce the measure t^*(Fubini-Study measure of $A_p^{[k]}$) on Λ and call it the "Fubini-Study measure".

- *Assumption 4.* Let k be as in *Assumption 3* and consider the sequence $j = 1, 2, \ldots, k$. Then, to each hyperplane in T_pA is associated a strictly decreasing sequence of non-empty Zariski closed subsets $\{\Lambda_j\}_{j=1}^k$ of Λ^0 consisting of those λ with the property that the local intersection number of S_λ and T_pA at p is at least j.

Later we will modify *Assumption 3 and 4* and the "Fubini-Study measure" in the following way:

1) We consider *Assumption 3 and 4* for various $k \in \mathbb{Z}_{>0}$. Namely, if we write Λ_k for such Λ satisfying the *Assumption 3 and 4* with positive integer k, we will consider families $\{S_\lambda\}_{\lambda \in \Lambda_k}$ for various $k \in \mathbb{Z}_{>0}$.

2) We consider a certain variation of the measure

$$t^*(\text{ “Fubini-Study measure” of } [A]^{(k)})$$

on the parameter space Λ. Namely, we will consider the variation of measures "interpolating" the Fubini-Study measure and the Dirac measure "perfectly concentrated" at some Λ_j associated to some hyperplane in T_pA. In particular, this

variation of measures contains a regular measures which strongly "concentrate" at some Λ_j (w.r.to some hyperplane of T_pA). Such a variation of measures will play an essential role in our proof of (1).

Let $k := S_\lambda \cdot D$ (the intersection number) and $m = g(S_\lambda)$ (the genus of S_λ, if S_λ is smooth). For $a \in A$ and a subset Z of A we set

$$Z + a = \{\zeta + a\,;\, \zeta \in Z\}\ .$$

This is the translate of Z by an element $a \in A$.

Definition 2.1. *Let $\phi_\lambda : A \to \mathrm{Hilb}^k(S_\lambda)$ denote the holomorphic map*

$$\phi_\lambda(a) = (D - a) \cap S_\lambda$$

repeated according to multiplicities. The **incidence divisor** *associated to D and S_λ is defined by*

$$D_\lambda := \phi_\lambda^*(\triangle)$$

where $\triangle$ is the incidence divisor of $\mathrm{Hilb}^k(S_\lambda)$.

Note that $\mathrm{Hilb}^k(S_\lambda)$ is a non-singular projective variety because S_λ is a curve (for instance, the elementary symmetric polynomials induces the well-known isomorphism $\mathrm{Hilb}^k(\mathbb{P}^1(\mathbb{C})) = \mathbb{P}^k(\mathbb{C})$). Let $f : \mathbb{C} \to A$ be an arbitrarily given holomorphic curve.

Definition 2.2. *(i) For each $\lambda \in \Lambda$, we define*

$$Y_\lambda := \text{the normalization of } \{(z,w) \in \mathbb{C} \times S_\lambda\,;\, w \in (D - f(z)) \cap S_\lambda\}\ ,$$

and an analytic covering map

$$\pi_\lambda\ :\ Y_\lambda \to \mathbb{C}\ ; \qquad (z,w) \mapsto z$$

of degree k and a holomorphic map

$$f_\lambda\ :\ Y_\lambda \to S_\lambda\ ; \qquad (z,w) \mapsto w\ .$$

(ii) The **Radon transformation** *of a given holomorphic curve $f : \mathbb{C} \to A$ with respect to $\{S_\lambda\}_{\lambda\in\Lambda}$ is defined to be the collection of holomorphic maps*

$$\{f_\lambda : Y_\lambda \to S_\lambda\}_{\lambda\in\Lambda}\ .$$

The basic tool to count the ramification of $f_\lambda : Y_\lambda \to S_\lambda$ is the Gauss map $\mathbb{C} \ni z \mapsto [f'(z)] \in \mathbb{P}(T_{f(z)}A)$. To identify the ramification counting function of $f_\lambda : Y_\lambda \to S_\lambda$ with a usual counting function with respect to a divisor, we introduce the notion of the Gaussian divisor:

Definition 2.2 (continued). *(i) The* **universal Gaussian divisor**

$$D_{\Lambda}^{[1]} \subset (\mathbb{P}(TA) \times \Lambda)$$

associated to D and the family $\{S_\lambda\}_{\lambda\in\Lambda}$ is defined by

$$D_{\Lambda}^{[1]} := \text{the Zariski closure in } \mathbb{P}(TA)\times\Lambda \text{ of}$$
$$\bigcup_{\substack{(a,\lambda)\in A\times\Lambda,\ D-a \text{ is non-singular} \\ \text{at the intersection } (D-a)\cap S_\lambda}} \left(\bigcup_{w\in(D-a)\cap S_\lambda} \underbrace{\mathbb{P}(T_w D)}_{\text{translated to } a} \times\{\lambda\}\right) .$$

(ii) The **Gaussian divisor** $D_{\lambda}^{[1]}$ *associated to S_λ is defined by*

$$D_{\lambda}^{[1]} := D_{\Lambda}^{[1]} \cap (\mathbb{P}(TA)\times\{\lambda\}) .$$

The Gaussian divisor $D_{\lambda}^{[1]}$ is a divisor of $\mathbb{P}(TA)$ such that $D_{\lambda}^{[1]} \cap \mathbb{P}(T_a A)$ consists of k hyperplanes in $\mathbb{P}(T_a A)$ for general $a \in A$. These hyperplanes are the collection of the translates to a of the projectivized tangent hyperplanes of $D-a$ at the intersection $(D-a)\cap S_\lambda$.

Let $k := S_\lambda \cdot D$. We have a k-valued holomorphic "map" $F_{D,\lambda} : A \to S_\lambda$ defined by $F_{D,\lambda}(a) = (D-a)\cap S_\lambda$ and the branched k-covering $A_\lambda \to A$ which transforms $F_{D,\lambda}$ into a single-valued holomorphic map from A_λ to S_λ. Clearly, A_λ is the "graph" in $A\times S_\lambda$ of the k-valued holomorphic map $F_{D,\lambda}$. We thus have a branched covering $\Pi_\lambda : A_\lambda \to A$. The incidence divisor D_λ in A is its branch divisor. We fix a non-zero holomorphic 1-form ω on A. Let ω_λ denote its restriction to the curve S_λ. Let $\widetilde{F}_{D,\lambda} : A_\lambda \to S_\lambda$ be the holomorphic map which uniformizes the k-valued holomorphic map $F_{D,\lambda} : A \to S_\lambda$. Let $\widetilde{D}(\omega_\lambda)_0$ be the pull-back of the canonical divisor $(\omega_\lambda)_0$ of S_λ via the holomorphic map $\widetilde{F}_{D,\lambda} : A_\lambda \to A$. Let $m := -(\text{Euler number of } S_\lambda) = 2g(S_\lambda) - 2 = \deg(\omega_\lambda)$. Then $\widetilde{D}(\omega_\lambda)_0$ defines a divisor $D(\omega_\lambda)_0$ on A which is of the form

$$D(\omega_\lambda)_0 = D_{\lambda,1} + \cdots + D_{\lambda,m}$$

where $D_{\lambda,1},\ldots,D_{\lambda,m}$ are translates of D in A.

Lemma 2.3 (Generalized Hurwitz formula). *The incidence divisor $D_\lambda \subset A$, the Gaussian divisor $D_\lambda \subset \mathbb{P}(TA)$ and the given divisor $D \subset A$ fit into the linear equivalence*

$$\pi^* D_\lambda - (\pi^* D_{\lambda,1} + \cdots + \pi^* D_{\lambda,m} + D_{\lambda}^{[1]}) = -k\, H^{[1]} . \tag{2}$$

Here, $\pi : \mathbb{P}(TA) \to A$ is the projection and $H^{[1]}$ is the relative hyperplane bundle of the projective vector bundle $\mathbb{P}(TA) \to A$, $m = g(S_\lambda)$ is the genus of the curve S_λ and $k = S_\lambda \cdot D$.

Proof. Using the local uniformization of the k-valued map $F_{D,\lambda} : A \to S_\lambda$ and the meromorphic vector field ω_λ^{-1} on S_λ, we can define a meromorphic section Σ_λ of the relative $\mathcal{O}(k)$ of $\mathbb{P}(TA) \cong A \times \mathbb{P}^{n-1}$ in the following way. To $a \in A$ we associate the

defining function $\Psi(a)$ of the union of k hyperplanes (counted with multiplicities) $\cup_{w\in(D-a)\cap S_\lambda}\mathbb{P}(T_wD)\subset\mathbb{P}(T_aA)$ determined by the equation

$$\Psi(a)\,(\underbrace{\omega_\lambda^{-1},\ldots,\omega_\lambda^{-1}}_{k-\text{times}})=1\ .$$

This equation determines uniquely the defining function $\Psi(a)$, if all intersections of $(D-a)$ and S_λ are transversal and $\omega_\lambda\neq 0$ at these points. Since the vector field ω_λ^{-1} has poles at points where $\omega_\lambda=0$, we have extra zeros $D(\omega_\lambda)_0$ (a divisor on A) in $(\Sigma_\lambda)_0$ corresponding to the zeros of ω_λ via the k-valued holomorphic map $F_{D,\lambda}:A\to S_\lambda$. This implies

$$(\Sigma_\lambda)_0=D_\lambda^{[1]}+\pi^*D(\omega_\lambda)_0\ .$$

On the other hand, if $(D-a)$ and S_λ has an intersection point w with multiplicity ≥ 2, the vector ω_λ^{-1} at T_wS_λ belongs to the kernel of the linear form defining the tangent space $T_w(D-a)$. This implies

$$(\Sigma_\lambda)_\infty=\pi^*D_\lambda\ .$$

Therefore we have the equality of divisors

$$(\Sigma_\lambda)_0-(\Sigma_\lambda)_\infty=D_\lambda^{[1]}+\pi^*D(\omega_\lambda)_0-\pi^*D_\lambda\ .$$

Thus we have the linear equivalence

$$k\,H^{[1]}=D_\lambda^{[1]}+(\pi^*D_{\lambda,1}+\cdots+\pi^*D_{\lambda,m})-\pi^*D_\lambda\ .$$

This completes the proof of Lemma 2.3. $\square$

We note that the right hand side $-k\,H^{[1]}$ of the generalized Hurwitz formula does not depend on D.

Lemma 2.4. *Counting the local intersection number of the projective jet lift $f^{[1]}:\mathbb{C}\to\mathbb{P}(TA)$ of f with the Gaussian divisor $D_\lambda^{[1]}\subset\mathbb{P}(TA)$ is equivalent to the local ramification counting of $f_\lambda:Y_\lambda\to S_\lambda$. Therefore, we have the identification between counting functions*

$$N_{f^{[1]},D_\lambda^{[1]}}(r)=N_{f_\lambda,\mathrm{Ram}}(r)$$

with a "slight exception". Here, the "slight exception" means the following : if $f'(z_0)=0$ holds, the "order" of the convergence (under the limit $w\to z_0$) of $[f'(w)]$ to $\bigcup_{w\in(D-f(z_0))\cap S_\lambda}\underbrace{\mathbb{P}(T_w(D-f(z_0)))}_{\text{translated to }f(z_0)}$ may be smaller than expected and thus we may have an inequality at the local level in the sense that the ramification index of f_λ is larger than the intersection number of $f^{[1]}$ and $D_\lambda^{[1]}$ at the point under consideration.

Proof. This is proved by local consideration and so it suffices to prove under the assumption that $f'(z)\neq 0$. We first consider the case that the intersection of

$f^{[1]} : \mathbb{C} \to \mathbb{P}(TA)$ occurs at $z = z_0 \in \mathbb{C}$ such that f touches D at $z = z_0$, i.e., $f(z_0) \in D$. Suppose that D is defined by an equation $h(z_1, \ldots, z_n) = 0$ where $z = (z_1, \ldots, z_n)$ is the local linear coordinate system of the Abelian variety. Let $f : \mathbb{C} \to A$ be a holomorphic curve given by $f(z) = (f_1(z), \ldots, f_n(z))$. Suppose that near $z = z_0$ we have

$$h(f_1(z), \ldots, f_n(z)) = O((z - z_0)^k) \ ,$$

i.e., the local intersection number of f and D at $z = z_0$ is k. Differentiating this implies

$$\frac{\partial h}{\partial z_1}(f(z))f_1'(z) + \cdots + \frac{\partial h}{\partial z_n}(f(z))f_n'(z) = O((z - z_0)^{k-1}) \ .$$

For each λ such that $f(z_0)+S_\lambda$ is tangent to D at $f(z_0)$ with multiplicity $\nu \geq 1$ ($\nu = 1$ means the transversal intersection), we define a ν-valued holomorphic "map" ψ_λ from a neighborhood of $f(z_0)$ to itself, by assigning w (in a neighborhood of $f(z_0)$) the nearest ν intersection points in $(w + S_\lambda) \cap D$. Choose a locally defined branch of the ν-valued holomorphic map ψ_λ and write it using the symbol $\psi_{\lambda,b}$. Then its image is contained in D and therefore $h(\psi_{\lambda,b}(f(z)) = 0$ holds for z sufficiently close to z_0. On the other hand, the tangent space at $\psi_{\lambda,b}(f(z))$ is given by the equation

$$\frac{\partial h}{\partial z_1}(\psi_{\lambda,b}(f(z)))\zeta_1 + \cdots + \frac{\partial h}{\partial z_n}(\psi_{\lambda,b}(f(z)))\zeta_n = 0 \ .$$

Write $\psi_{\lambda,b}(f(z)) = f(z)+\phi_{\lambda,b}(z) = (f_1(z)+\phi_{\lambda,b,1}(z), \ldots, f_n(z)+\phi_{\lambda,b,n}(z))$. Then, from these two equations, we have

$$\begin{aligned}
&\frac{\partial h}{\partial z_1}(\psi_{\lambda,b}(f(z)))f_1'(z) + \cdots + \frac{\partial h}{\partial z_n}(\psi_{\lambda,b}(f(z)))f_n'(z) \\
&= \frac{\partial h}{\partial z_1}(f(z) + \phi_{\lambda,b}(z))(f_1'(z)) + \cdots + \frac{\partial h}{\partial z_n}(f(z) + \phi_{\lambda,b}(f(z))(f_n'(z)) \\
&= -\frac{\partial h}{\partial z_1}(f(z))f_1'(z)) + \cdots + \frac{\partial h}{\partial z_n}(f(z))f_n'(z)) + O((z - z_0)^{\frac{k-1}{\nu}}) \\
&= O((z - z_0)^{\frac{k-1}{\nu}}) \ .
\end{aligned}$$

We have ν such estimates corresponding to the nearest ν intersection points of S_λ and $D - f(z_0)$. This way, we measure the "distance" between $[f'(z)] \in \mathbb{P}(T_{f(z)}A)$ and $\cup_{w\in(D-f(z))\cap S_\lambda} \underbrace{\mathbb{P}(T_w(D - f(z)))}_{\text{translated to } f(z)}$ in terms of the order of divisibility w.r.to $(z - z_0)$. Therefore, the local intersection number of $f^{[1]}$ and $D_\lambda^{[1]}$ at $z = z_0$ and the local ramification index of f_λ at $z = z_0$ are the same. We note that this argument does not depend on the multiplicity ν of the intersection of $(D - f(z_0)) \cap S_\lambda$ at the neutral point p.

Next, we consider the case that the intersection of $f^{[1]} : \mathbb{C} \to \mathbb{P}(TA)$ occurs at $z = z_0 \in \mathbb{C}$ such that $f(z_0) \notin D$. We get the same local conclusion in this case also. It follows from the definition of $D_\lambda^{[1]}$ that we have only to work at such an intersection point of $D - f(z_0)$ and S_λ with multiplicity ≥ 2. We want to measure the "distance" between $[f'(z)] \in \mathbb{P}(T_{f(z)}A)$ and $\cup_{w\in(D-f(z))\cap S_\lambda} \mathbb{P}(T_w(D - f(z)))$

in terms of the order of divisibility w.r.to $(z - z_0)$. We can do this even if the location of $f(z_0)$ is far from D. What is important in this consideration is the jet of f at $z = z_0$. Indeed, the jet is parallel translated to any place in a portion of S_λ close to the intersection point with $D - f(z_0)$ by using the group structure of the Abelian variety A. $\square$

In the proof of the following Lemma 2.5, we will make a certain observation on the counting functions of the ramification of $\pi_\lambda : Y_\lambda \to \mathbb{C}$ and the intersection of f and D_λ. Namely, we will introduce the variation of the Fubini-Study measure of Λ (the variation parametrized by t, positive real numbers in the definition of the counting function) which was mentioned after *Assumption 1-4* and show that the "concentration" has an effect on the "averaging" procedure of the difference of two counting functions $N_{\pi_\lambda,\mathrm{Ram}}(r) - N_{f,D_\lambda}(r)$. Namely, we are interested in the "response" of the average[3] of two counting functions $N_{\pi_\lambda,\mathrm{Ram}}(r)$ and $N_{f,D_\lambda}(r)$, under the singular perturbation ("concentration") of the measures of Λ. The reason why we are interested in the "response" is the following. The ramification counting function $N_{\pi_\lambda,\mathrm{Ram}}(r)$ does not count the order of tangency of f to D "honestly", even if we introduce such Λ with large k in *Assumption 3 and 4*. On the other hand, provided we are working under the existence of an upper bound for r, the counting function $N_{f,D_\lambda}(r)$ counts the tangency of f to D "honestly", if we introduce such Λ with sufficiently large k (depending on the upper bound of r) in *Assumption 3 and 4*.

We now introduce the "variation" of measures on Λ in the following way. To do so, we take, as an example, the averaging procedure of $N_{f,D_\lambda}(r)$ over Λ with a variation of measures. Suppose that f and D intersects with multiplicity ≥ 2 at $z = z_0$. If S_λ is tangent to $D - f(z_0)$ at the neutral point p with sufficiently large intersection multiplicity, this multiple intersection point $f(z_0)$ of f and D is counted in the counting function $N_{f,D_\lambda}(r)$. This is the reason why we take $N_{f,D_\lambda}(r)$ as an example. We recall the definition of the counting function

$$N_{f,D_\lambda}(r) = \int_0^r \frac{n_{f,D_\lambda}(t) - n_{f,D_\lambda}(0)}{t} dt + n_{f,D}(0) \log r \ .$$

Fix $r > 0$. Let $z_{01}, z_{02}, \ldots, z_{0k} \in \mathbb{C}$ be the places z of $\mathbb{C}$ where there exist multiple intersections in $(D - f(z)) \cap S_\lambda$ at the neutral point p. We introduce the following slight modification (if necessary) in counting the intersection of f and D_λ. Namely, if we have several multiple intersections of f and D on the circle $\partial\mathbb{C}(t) = \{z \in \mathbb{C}\,; |z| = t\}$, we perturb these points so that these points are separated by the distance from the origin, i.e., if z_{01} and z_{02} are such points, i.e., $|z_{01}| = |z_{02}| = t$, then we will count these multiple intersections (in the definition of $N_{f,D_\lambda}(r)$) as if $|z_{01}|$ and $|z_{02}|$ are very close to t and different. Of course we modify only slightly so that the difference of the modified counting function and the original counting function is negligible in the sense that the difference is at most $\varepsilon T_{f,D}(r)$ where $\varepsilon > 0$ is an arbitrarily given small number. We are now ready to introduce the variation of measures on Λ. Let k be a positive number which is sufficiently large compared to the largest intersection multiplicity of $f_{\mathbb{C}(r)} : \mathbb{C}(r) \to A$ and D. We then introduce the measured space $(\Lambda, d\lambda_t)$ $(t \leq r)$ by the following rule. For t such that $0 < t < |z_{01}|$, $d\lambda_t$ is a usual Fubini-Study probability measure. Suppose

---

[3] Of [illegible]

that D is non-singular at the intersection point of f and D at $f(z_{01})$. For t such that $|z_{0|}| \leq t < |z_{02}|$, we define $d\lambda_t$ as a regular probability measure "concentrated" at Λ_{j_1} corresponding to the hyperplane in T_pA which is the parallel translate to the neutral point p of the tangent plane of D at the multiple intersection point $f(z_{01})$ with f (here, we assume that D is non-singular at $f(z_{01})$), where j is the intersection multiplicity of f and D at $z = z_{01}$. If $f(z_0)$ belongs to the singular locus of D, we need a slight modification. However, the modification is easy, because D is assumed to have at worst normal crossings. Next, suppose that D is non-singular at the intersection point of f and D at $f(z_{02})$. For t such that $|z_{02}| \leq t < |z_{03}|$, we define $d\lambda_t$ as a regular probability measure "concentrated" at Λ_{j_2} corresponding to the hyperplane in T_pA which is the parallel translate to the neutral point p of the tangent plane of D at the multiple intersection point $f(z_{02})$ with f (here, we assume that D is non-singular at $f(z_{02})$), where j_2 is the intersection multiplicity of f and D at $z = z_{02}$. Again the necessaey modification when D is singular at $f(z_{02})$ is easy. Repeating this procedure, we get a variation of probability measures $d\lambda_t$ $(0 < t < r)$ and get the measured space $(\Lambda, d\lambda_t)$ parametrized by t $(0 < t < r)$. We then define the "average of the counting function $N_{f,D_\lambda}(r)$ over the variation of probability spaces $(\Lambda, d\lambda_t)$", denoted by $N^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r)$, as follows:

$$N^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r) := \int_0^r \frac{dt}{t} \int_{(\Lambda,d\lambda_t)} (n_{f,D_\lambda}(t) - n_{f,D_\lambda}(0)) d\lambda_t + n_{f,D}(0) \log r \ .$$

For other Nevanlinna theoretic functions, we define the corresponding "average" over the variation of probability spaces (Λ, λ_t) in the following way:

$$T^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r) := \int_0^r \frac{dt}{t} \int_{(\Lambda,d\lambda_t)} d\lambda_t \int_{\mathbb{C}(t)} f^* c_1(\mathcal{O}_A(D_\lambda)) \ ,$$

$$m^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r) := \int_0^{2\pi} \frac{d\theta}{2\pi} \int_{(\Lambda,d\lambda_r)} \log \frac{1}{\mathrm{dist}_{\mathrm{Euc}}(f(re^{i\theta}), D_\lambda)} \ ,$$

The response under this variation of probability spaces is summarized in the following:

Lemma 2.5. *For any variation of probability spaces $(\Lambda, d\lambda_t)$ $(0 < t < r)$ as above, we have*

$$m^{(\Lambda,d\lambda_t)}_{f_\lambda,p}(r) \ = \ m_{f,D}(r) + O(1) \ . \tag{3}$$

On the other hand, let ε be any positive number. Then there exists a variation of probability measures $(\Lambda, d\lambda_t)$ $(0 < t < r)$ with sufficiently strong "concentration" such that the "averaging" over $(\Lambda, d\lambda_t)$ satisfies the following estimates :

$$N^{(\Lambda,d\lambda_t)}_{f_\lambda,\mathrm{Ram}}(r) - N^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r) \ = 0 \ , \tag{4}$$

$$N^{(\Lambda,d\lambda_t)}_{\pi_\lambda,\mathrm{Ram}}(r) - N^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r) \ \leq \ -N^1_{f,D}(r) + \varepsilon\, T_{f,D}(r) \ . \tag{5}$$

Here, p in (3) is the neutral point (see the beginning of §2), the $O(1)$ term in (3) is independent of the variation $(\Lambda, d\lambda_t)$ and $N^1_{f,D}(r)$ in (4) is the residual counting

function which counts all intersections with multiplicity 1*. Moreover, the equality (4) holds "with slight exceptions" (i.e., in the local level the equality* $=$ *may be inequality* $\geq$ *at* z *such that* $f'(z) = 0$*) in the same sense as in Lemma 2.4.*

Proof. Since $p \in S_\lambda$ for all $\lambda \in \Lambda$ (p is the neutral point of A) and f approximates D if and only if f_λ approximates p, we have the averaging formula (3). Indeed, it is easy to prove (3) by combining the argument in Lemma 2.4 and [K3, Lemma 2.2]. Indeed, [K3, Lemma 2.2] implies that the proof of (3) is based on the estimate of the functions on the projective space which has logarithmic poles along the hyperplane at infinity and Lemma 2.4 implies that the estimate is stable for large k with bound (recall that, in the proof of Lemma 2.4, we have the multiplicity ν in the denominator of the exponent of $(z - z_0)$ and we have ν such estimates). The estimate "with slight exception" is a consequence of Lemma 2.4. To prove (5), we choose the family of curves $\{S_\lambda\}_{\lambda\in\Lambda}$ with k in *Assumption 3 and 4* sufficiently large compared to the maximum multiplicity of f and D in $\mathbb{C}(r)$. We can then choose a variation of probability spaces (Λ, λ_t) in such a way that it picks up all multiple intersections of f and D in $\mathbb{C}(r)$ just like in the above argument. If the intersection multiplicity of S_λ and $D - f(z_0)$ at p is sufficiently large compared to the intersection multiplicity of f and D at $z = z_0$, then the intersection of f and D_λ is equivalent to that of f and D at $z = z_0$. In this situation, the counting function $N_{f,D_\lambda}(r)$ counts all multiple intersections of multiplicity ν (of f and D in $\mathbb{C}(r)$) as those of multiplicity $\nu - 1$. On the other hand, the counting function $N_{\pi_\lambda,\mathrm{Ram}}(r)$ does not count the multiplicity of f and D. Indeed, the k-covering map $\pi_\lambda : Y_\lambda \to \mathbb{C}$ is obtained by the pull back (via $f : \mathbb{C} \to A$) of the branched k-covering $A_\lambda \to A$ uniformizing the k-valued holomorphic map $A \to S_\lambda$ defined by $A \ni a \mapsto (D - a) \cap S_\lambda$ (cf. the proof of Lemma 2.3). Therefore, given small $\varepsilon > 0$, if the "concentration" to the sequence of Λ_j's which appear corresponding to the sequence of multiple intersection points (of f and D in $\mathbb{C}(r)$) is sufficiently strong, then we have the estimate of the form of (5). □

The following corollary is a consequence of (3) in Lemma 2.5 (computation in the proof of Lemma 2.4), or more generally, the invariance of the proximity functions under "measure concentration".

Corollary 2.6. *The First Main Theorem* $m_{f,D_\lambda}(r) + N_{f,D_\lambda}(r) = T_{f,D_\lambda}(r) + O(1)$ *remains to be true after taking the "average" over the variation of probability spaces* $(\Lambda, d\lambda_t)$*, i.e., we have*

$$m^{(L,d\lambda_t)}_{f,D_\lambda}(r) + N^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r) = T^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r) + O(1)\ .$$

To sum up, we have introduced the Gaussian divisor $D^{[1]}_\lambda$ and the incidence divisor D_λ to interpret the ramification counting functions $N_{f_\lambda,\mathrm{Ram}}(r)$ and $N_{\pi_\lambda,\mathrm{Ram}}(r)$. To get information which is independent of λ, we have to take the "average" over Λ against some probability measure. To get such information, we have introduced the "variation of probability spaces" $(\Lambda, d\lambda_t)$ and look at the difference of the response under this "variation of probability measures" of Λ. Namely, the "concentration" of probability measures at subvarieties Λ_j's of Λ corresponding to high degree of tangents of S_λ and D. As a result, the difference $N_{f_\lambda,\mathrm{Ram}}(r) - N_{f^{[1]},D^{[1]}_\lambda}(r)$ is essentially stable under the "variation of probability spaces" $(\Lambda, d\lambda_t)$. On the contrary,

the quantity $-N^1_{f,D}(r)$ stemming from the multiple intersection (of f and D in $\mathbb{C}(r)$), which is only a small portion of the difference $N_{\pi_\lambda,\mathrm{Ram}}(r) - N_{f,D_\lambda}(r)$ w.r.to the average against the usual Fubini-Study measure, becomes "dominating" under the average against the variation of the probability space $(\Lambda, d\lambda_t)$ with sufficiently strong "concentration".

There is a natural way[4] of producing the variation of probability spaces $(\Lambda, d\lambda_t)$ which is useful in the proof of Lemma 2.5. We introduce the notion of the "combinatorial blow up" of Λ. We "expand" the parameter space Λ in a combinatorial way without changing the family of curves itself. Here, expanding the parameter space Λ in a combinatorial way *e.t.c.* means the following. Suppose that λ is a positive number valued coordinate function of Λ (for instance, some angle parameter naturally defined on Λ by using the Hopf fibration from odd dimensional sphere to a complex projective space). If we consider the totality of all decompositions of λ into N positive numbers $\lambda = \lambda_1 + \cdots + \Lambda_N$, we can increase the degree of freedom in the coordinate[5]. We call this the "combinatorial blow up". We then introduce the natural probability measure on the combinatorial blow up. This has the effect that the neighborhood of the fixed radius of a point becomes relatively smaller, if N becomes larger. Applying the combinatorial blow up to Λ, we can increase its dimension without increasing the family itself (i.e., this increases only the parameter space without increasing the family). The tuple $(\lambda_1, \ldots, \lambda_N)$ represents the same S_λ for $\forall(\lambda^1_N, \ldots, \lambda^N_N)$ if $\lambda = \lambda_N = \lambda^1_N + \cdots + \lambda^N_N$.

The following Lemma 2.7 is a direct consequence of Lemma 2.5 and explains the reason why one can reduce the truncation level to 1 in the Second Main Theorem for holomorphic curves into Abelian varieties. Let $N^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r)$, *etc*, be the same as in Lemma 2.5.

Lemma 2.7. *Let A be an Abelian variety, D any ample reduced divisor, $\varepsilon > 0$ any positive number and $f : \mathbb{C} \to A$ any holomorphic curve such that $f(\mathbb{C}) \not\subset \mathrm{Supp}(D)$. Then there exists a variation of probability measures $(\Lambda, d\lambda_t)$ $(0 < t < r)$ with sufficiently strong "concentration" such that the "averaging" over $(\Lambda, d\lambda_t)$ satisfies the following estimates :*

$$\begin{aligned} &N^{(\Lambda,d\lambda_t)}_{f_\lambda,\mathrm{Ram}}(r) - N^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r) + N^{(\Lambda,d\lambda_t)}_{f,D_{\lambda_N}}(r) - N^{(\Lambda,d\lambda_t)}_{\pi_\lambda,\mathrm{Ram}}(r) \\ &\quad \geq N^1_{f,D}(r) - \varepsilon\, T_{f,D}(r) /\!/ \,. \end{aligned}$$

We are ready to prove our main result (1).

Let $f : \mathbb{C} \to A$ be a holomorphic curve into an Abelian variety A and D an ample reduced divisor in A (for simplicity, we assume that D has at worst normal crossings).

From here on, we consider the family of algebraically equivalent algebraic curves $\{S_\lambda\}_{\lambda\in\Lambda}$ satisfying the *Assumptions 1-4* stated at the beginning of this §2. In particular, the parameter space Λ is equipped with the variation of the probability measures $d\lambda_t$ so that Lemma 2.5 and 2.7 hold. In particular, we consider the variation of the probability spaces $(\Lambda, d\lambda_t)$ where the measure $d\lambda_t$ "concentrates"

[4] We use the stratification of Λ into Λ_j's w.r.to a fixed hyperplane of T_pA in *Assumption 4*.

[5] This is something similar to the blow up, because we replace a point λ_0 on a λ-line by the

at Λ_j's sufficiently strongly, so that we have the estimates in (3), (4), (5) in Lemma 2.5 and the estimate in Lemma 2.7.

We start by applying the Second Main Theorem to the family of holomorphic maps $f_\lambda : Y_\lambda \to \mathbb{C}$ and "average" the result over the variation of probability spaces (Λ, λ_t).

The Second Main Theorem for holomorphic maps from a finite analytic covering space of $\mathbb{C}$ to a compact Riemann surface is available in, for instance, [N]. Let p denote the neutral point of the abelian variety A. Then for all $\lambda \in \Lambda$ the curve S_λ contains p. Now the Second Main Theorem for $f_\lambda : Y_\lambda \to \mathbb{C}$ is stated as follows:

$$(8) \quad m_{f_\lambda,p}(r) + T_{f,K_{S_\lambda}}(r) + N_{f_\lambda,\mathrm{Ram}}(r) - N_{\pi_\lambda,\mathrm{Ram}}(r) \le O(\log T_{f,D}(r) + \log r) /\!/ \,.$$

The right hand side is bounded above by the quantity of the form $k \log T_{f,D}(r)$, where $k = S_\lambda \cdot D = [Y_\lambda : \mathbb{C}]$ is the covering degree of $\pi_\lambda : Y_\lambda \to \mathbb{C}$.

Since $p \in S_\lambda$ for all $\lambda \in \Lambda$ and f approximates D if and only if f_λ approximates p, (3) in Lemma 2.5 implies the averaging formula

$$m_{f,D}(r) = m^{(\Lambda,d\lambda_t)}_{f_\lambda,p}(r)d\lambda + O(1) \ ,$$

where the $O(1)$-term is independent of f.

Next we would like to apply the above Second Main Theorem to f_λ and average the result over the variation of probability spaces (Λ, λ_t) in the sense of Lemma 2.5. Here, the existence of exceptional intervals in the Second Main Theorem may cause difficulty. However, the size of the exceptional intervals implicit in the symbol $/\!/_\varepsilon$ is controlled by the Borel Lemma, whose "family version" is formulated and proved in [K2, (2.6)]. This implies that we can average the Second Main Theorems for f_λ's without worrying about the individual exceptional intervals. Moreover, the Second Main Theorem is a consequence of Nevanlinna's lemma on logarithmic derivative (see, for instance [NoO] and [L2]). The constant term depends on λ and might cause problem in the course of averaging. This constant term in turn stems from the constant term in the First Main Theorem for f_λ and is of the form $\log|c_\lambda|$, where c_λ is determined by the behavior of f_λ at $z = 0$. As f is fixed and f_λ is defined algebraically using the intersection $(D - f(z)) \cap S_\lambda$ $(\lambda \in \Lambda)$, c_λ depends algebraically on λ and therefore the integral of $\log|c_\lambda|$ over Λ is finite whose value depends only on the behavior of f at $z = 0$ and $\{S_\lambda\}_{\lambda\in\Lambda}$.

We may now apply the Second Main Theorem (8) to each f_λ and average the result over (Λ, λ_t) in the sense of Lemma 2.5. Using the concavity of the logarithm, we have

$$\begin{aligned} &m_{f,D}(r) = m^{(\Lambda,d\lambda_t)}_{f_\lambda,p}(r)d\lambda + O(1) \\ &\le -T^{(\Lambda,d\lambda_t)}_{f_\lambda,K_{S_\lambda}}(r) - (N^{(\Lambda,d\lambda_t)}_{f_\lambda,\mathrm{Ram}}(r) - N^{(L,d\lambda_t)}_{\pi_\lambda,\mathrm{Ram}}(r)) + O(\log T^{(\Lambda,d\lambda_t)}_{f_\lambda,p}(r) + \log r) /\!/ \,. \end{aligned}$$

It follows from the definition that the averaging formula

$$T_{f,D}(r) = T^{(\Lambda,d\lambda_t)}_{f_\lambda,p}(r)d\lambda + O(1)$$

holds, where $O(1)$-term does not depend on f. Let $\deg K_{S_\lambda} = m$ and $D_{\lambda,1}, \ldots, D_{\lambda,m}$ be as in Lemma 2.3. Then, we have

$$T^{(\Lambda,d\lambda_t)}(r)d\lambda = T_{f,D_{\lambda,1}+\cdots+D_{\lambda,m}}(r) + O(1)$$

Applying Lemma 2.7, we get

$$\begin{aligned} m_{f,D}(r) \le & \{-T^{(\Lambda,d\lambda_t)}_{f,D_{\lambda,1}+\cdots+D_{\lambda,m}}(r) - N^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r) + N^{(\Lambda,d\lambda_t)}_{f,D_\lambda}(r)\}\} \\ & - \{(N^{(\Lambda,d\lambda_t)}_{f_\lambda,\mathrm{Ram}}(r) - N^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r)) - (N^{(\Lambda,d\lambda_t)}_{\pi_\lambda,\mathrm{Ram}}(r) - N^{(\Lambda,\lambda_t)}_{f,D_\lambda}(r))\} \\ & + O(\log T_{f,D}(r) + \log r) + O(1)/\!/ \\ \le & \{-T^{(\Lambda,d\lambda_t)}_{f,D_{\lambda,1}+\cdots+D_{\lambda,m}}(r) - N_{f^{[1]},D^{[1]}_\lambda}(r) + N_{f,D_\lambda}(r)\} \\ & - N^1_{f,D}(r) + \varepsilon\, T_{f,D}(r) + O(\log T_{f,D}(r) + \log r) + O(1)/\!/ \ . \end{aligned}$$

We now apply the First Main Theorem (Corollary 2.6) to replace the counting function $N_{f^{[1]},D^{[1]}_\lambda}(r)$ (resp. $N_{f,D_\lambda}(r)$) by the difference of the height function and the proximity function $T_{f^{[1]},D^{[1]}_\lambda}(r) - m_{f^{[1]},D^{[1]}_\lambda}(r)$ (resp. $T_{f^{[1]},\pi^*D_\lambda}(r) - m_{f^{[1]},\pi^*D_\lambda}(r)$):

$$\begin{aligned} m_{f,D}(r) + N^1_{f,D}(r) = & \{-T^{(\Lambda,d\lambda_t)}_{f^{[1]},\pi^*D_{\lambda,1}+\cdots+\pi^*D_{\lambda,m}}(r) - T^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r) + T^{(\Lambda,d\lambda_t)}_{f^{[1]},\pi^*D_\lambda}(r)\} \\ & + m^{(\Lambda,\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r) - m^{(\Lambda,d\lambda_t)}_{f^{[1]},\pi^*D_\lambda}(r) \\ & + \varepsilon\, T_{f,D}(r) + O(\log T_{f,D}(r) + \log r) + O(1)/\!/ \ . \end{aligned}$$

As $\{D_\lambda\}_{\lambda\in\Lambda}$ is free from base locus, the computation in the proof of Lemma 2.5 implies

$$\int_{\lambda\in\Lambda} m_{f^{[1]},\pi^*D_\lambda}(r)d\lambda = O(1) \ .$$

Hence

$$\begin{aligned} m_{f,D}(r) + N^1_{f,D}(r) \ \le & \ T^{(\Lambda,d\lambda_t)}_{f^{[1]},-m\,\pi^*D-D^{[1]}_\lambda+\pi^*D_\lambda}(r) + m^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda} \\ & + \varepsilon\, T_{f,D}(r) + O(\log T_{f,D}(r) + \log r) + O(1)/\!/ \ . \end{aligned}$$

Applying the Generalized Hurwitz formula (2) in Lemma 2.3, we have

$$\begin{aligned} m_{f,D}(r) + N^1_{f,D}(r) \ \le & \ -k\, T_{f^{[1]},H^{[1]}}(r) + m^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r) \\ & + \varepsilon\, T_{f,D}(r) + O(\log T_{f,D}(r) + \log r) + O(1)/\!/ \ . \end{aligned}$$

It follows from the definition of the Gaussian divisor that

$$\begin{aligned} \mathrm{Bs}\,\{D^{[1]}_\lambda\}_{\lambda\in\Lambda} & = \text{Zariski closure of } \mathbb{P}(TD_{\mathrm{reg}}) \\ & =: D^{[1]} \quad (\subset \mathbb{P}(TA)) \ . \end{aligned}$$

This implies that, averaging $m_{f^{[1]},D^{[1]}_\lambda}(r)$ over Λ, we have a main contribution of the base locus $D^{[1]}$ as well as that coming from those curves S_λ which are tangent to $D - w$ ($w \in D$) at the origin. The latter is not larger than $\varepsilon\, T_{f,D}(r)$ for any positive ε. Namely we have the following Lemma.

Lemma 2.8. *We have*

$$m^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r)d\lambda \ \leq \ m_{f^{[1]},D^{[1]}}(r) + \varepsilon\, T_{f,D}(r)\ .$$

Proof. We write the average $m^{(\Lambda,d\lambda_t)}_{f^{[1]},D^{[1]}_\lambda}(r)$ as

$$\int_0^{2\pi}\left(\int_{(\Lambda,\lambda_r)} \log^+ \frac{1}{\operatorname{dist}_{\mathrm{Euc}}(f^{[1]}(re^{i\theta}), D^{[1]}_\lambda)} d\lambda_r\right)\frac{d\theta}{2\pi}\ .$$

The main part of the integral over Λ consists of the contribution of the base locus $D^{[1]}$, i.e.,

$$\log^+ \frac{1}{\operatorname{dist}_{\mathrm{Euc}}(f^{[1]}(re^{i\theta}), D^{[1]})}$$

and possibly the contribution from those curves S_λ which are tangent to $D-w$ ($w\in D$) at the neutral point. To see the reason why we have to take the latter contribution into account, we take an arbitrary $w \in D$. Let $\lambda \in \Lambda$ be such that S_λ is tangent to $D-w$ at the neutral point. Then the Gaussian divisor $D^{[1]}_\lambda$ contains $D^{[1]}$ with multiplicity ≥ 2. So, if f approximates $w\in D$, we have to take special care of those S_λ's tangent to $D-w$ at the neutral point. The latter contribution is not larger than the product of the volume of the ε-tubular neighborhood of the subvariety Λ_w of Λ parameterizing those curves and the proximity function of $f^{[1]}$ to the Gaussian divisor $D^{[1]}_\lambda \subset \mathbb{P}(TA)$ ($\lambda \in \Lambda_w$). So, if the measure under consideration is "concentrated", the main contribution comes from those λ such that the corresponding S_λ intersects $D-w$ ($w\in D$) at the neutral point with high multiplicity. However, it follows from the computation in Lemma 2.4 that the contribution to the proximity function $m_{f^{[1]},D^{[1]}_\lambda}(r)$ is independent of the multiplicity of S_λ and $D-w$. Therefore, the main part of the left hand side of Lemma 2.6 is $m_{f^{[1]},D^{[1]}}(r)$. The error of this approximation comes from the portion of $D^{[1]}_\lambda$ close to the base locus $D^{[1]}$. On the other hand, the approximation of $f^{[1]}$ to $D^{[1]}_\lambda$ is dominated by the height function $T_{f^{[1]},D^{[1]}_\lambda}(r)$ and this height function is bounded from above and below by the height function $T_{f,D}(r)$ in the sense that there exist positive constants C_1 and C_2 such that

$$C_1\, T_{f,D}(r) \leq T_{f^{[1]},D^{[1]}_\lambda}(r) \leq C_2\, T_{f,D}(r)$$

holds. Therefore the the latter contribution is not larger than $\varepsilon\, T_{f,D}(r)$ for any positive number ε in the asymptotic sense when $r \to \infty$. $\square$

Summing up the above argument, we have

Proposition 2.9. *Let A be an Abelian variety, D a reduced divisor and $\{S_\lambda\}_{\lambda\in\Lambda}$ an algebraic family of algebraically equivalent curves in A all passing through the neutral point p such that general S_λ is a non-singular curve and the natural rational map $\Lambda \cdots \to \mathbb{P}(T_pA)$ is surjective. Let $H^{[1]}$ (resp. $D^{[1]}_\lambda$) denote the relative hyperplane bundle of $\mathbb{P}(TA)$ (resp. the Gaussian divisor with respect to $\{S_\lambda\}_{\lambda\in\Lambda}$). Let ε be any positive number. Then we have*

$$m_{f,D}(r) + N^1_{f,D}(r) \ \leq \ -k\, T_{f^{[1]},H^{[1]}}(r) + m_{f^{[1]},D^{[1]}}(r)$$
$$+\, \varepsilon T_{f,D}(r) + O(\log T_{f,D}(r) + \log r) + O(1)\,//$$

So far we worked over A and $A^{[1]} = \mathbb{P}(TA)$. However, we are able to perform all above arguments over the cone $A^{(1)} := \mathbb{P}(TA \oplus \mathcal{O}_A)$. Here we list all definitions which should be modified if we replace $A^{[1]}$ with $A^{(1)}$.
(i) The Gaussian divisor $D_\lambda^{[1]}$ should be replaced with its cone $D_\lambda^{(1)}$ (i.e., $A^{[1]}$ is the hyperplane at infinity in $A^{(1)}$ and $D^{(1)}$ is defined to be the cone over $D_\lambda^{[1]}$ in $A^{(1)}$).
(ii) The projective jet lift $f^{[1]}$ should be replaced with its cone $f^{(1)}(z) := [f'(z) \oplus 1_{f(z)}]$, where 1_a ($a \in A$) is the value of the generator of $H^0(A, \mathcal{O}_A)$ at a.
(iii) The relative hyperplane bundle $H^{[1]}$ of the projective vector bundle $A^{[1]} \to A$ should be replaced with the relative hyperplane bundle $H^{(1)}$ of the cone $A^{(1)} \to A$.
(iv) $D^{[1]}$ should be replaced with its cone $D^{(1)}$.

All the other definitions (D_λ, Y_λ, $\pi_\lambda : Y_\lambda \to \mathbb{C}$ and $f_\lambda : Y_\lambda \to S_\lambda$) are the same.

The Generalized Hurwitz formula (2) remains true, if we replace $H^{[1]}$ (resp. $\pi : A^{[1]} \to A$ and $D_\lambda^{[1]}$) with $H^{(1)}$ (resp. their cone $\pi : A^{(1)} \to A$ and $D_\lambda^{(1)}$).

The crucial Lemmas 2.5 and 2.7 continue to hold, if we replace $f^{[1]}$ (resp. $D_\lambda^{[1]}$) with its "cone" $f^{(1)}$ (resp. $D_\lambda^{(1)}$).

As a result we have

Proposition 2.10. *Let A be an Abelian variety, D a reduced divisor and $\{S_\lambda\}_{\lambda\in\Lambda}$ an algebraic family of algebraically equivalent curves in A all passing through the neutral point p such that general S_λ is a non-singular curve and the natural rational map $\Lambda \cdots \to \mathbb{P}(T_pA)$ is surjective. Let $H^{(1)}$ (resp. $D_\lambda^{(1)}$) denote the relative hyperplane bundle of $\mathbb{P}(TA \oplus \mathcal{O}_A)$ (resp. the cone of the Gaussian divisor with respect to $\{S_\lambda\}_{\lambda\in\Lambda}$). Let ε be any positive number. Then we have*

$$\begin{aligned} m_{f,D}(r) + N^1_{f,D}(r) \ \le\ & -k\, T_{f^{(1)},H^{(1)}}(r) + m_{f^{(1)},D^{(1)}}(r) \\ & + \varepsilon\, T_{f,D}(r) + O(\log T_{f,D}(r) + \log r) + O(1) /\!/_{\varepsilon} \ . \end{aligned}$$

Therefore, the proof of (1) is reduced to the proof of the estimate of type

$$m_{f^{(1)},D^{(1)}}(r) \le \varepsilon\, T_{f,D}(r) \ .$$

In §3, we solve this problem by using a geometric version of Nevanlinna's Lemma on Logarithmic Derivative ([Y1], [K3]).

3. Lemma on logarithmic derivative.

It follows from Proposition 2.10 that the proof of (1) is reduced to the estimates of $T_{f^{(1)},H^{(1)}}(r)$ and $m_{f^{(1)},D^{(1)}}(r)$. For this purpose we consider the sequence of successive projective completion of tangent bundles

$$\begin{aligned} &A^{(0)} = A,\ A^{(1)} = \mathbb{P}(TA \oplus \mathcal{O}_A),\ A^{(2)} = \mathbb{P}(TA^{(1)} \oplus \mathcal{O}_{A^{(1)}}), \ldots, \\ &A^{(i)} = \mathbb{P}(TA^{(i-1)} \oplus \mathcal{O}_{A^{(i-1)}}), \ldots\ . \end{aligned}$$

A geometric version of the Lemma on logarithmic derivative ([Y1], [V2] and [K1]) yields the estimates for $T_{f^{(1)},H^{(1)}}(r)$ and $m_{f^{(1)},D^{(1)}}(r)$. Here we recall this. Let X be a smooth projective variety. Write $X^{(i)}$ for the i-th successive projective completion of tangent bundles. Let Z a subscheme given by $Z = V(f_1, \ldots, f_k)$. Then $Z^{(i)}$ is a subvariety of $X^{(i)}$ defined by the Zariski closure of $Z^{(i)}_{\mathrm{reg}}$ in $A^{(i)}$. By the symbol ∞, we denote the divisor at infinity in the projective completion of a vector bundle. We are now ready to state a geometric version of Lemma on logarithmic derivative.

Lemma 3.1. *Let X be a smooth projective variety, Z any subscheme and $E \to X$ any ample line bundle. Let $f : \mathbb{C} \to X$ be an arbitrary holomorphic curve such that $f(\mathbb{C}) \not\subset \mathrm{Supp}(Z)$. Then we have*

$$\begin{cases} m_{f^{(i)},Z^{(i)}}(r) \le m_{f,Z}(r) + O(\log^+(r\,T_{f,E}(r)))/\!/\ , \\ m_{f^{(i)},\infty}(r) \le O(\log^+(r\,T_{f,E}(r)))/\!/\ . \end{cases}$$

The advantage of this formulation is that the original form

$$m_{f'/f,\infty}(r) \le O(\log^+(rT_f(r)))/\!/$$

of Nevanlinna's lemma on logarithmic derivative for $f : \mathbb{C} \to \mathbb{P}^1(\mathbb{C})$ splits into two inequalities each of which has its clear geometric meaning. For the proof, we refer to [Y1], [V2] and [K1,3].

We return to our situation. We apply Lemma 3.1 to the holomorphic curve $f : \mathbb{C} \to A$. We put $X = A$, $Z = D$ in Lemma 3.1. We fix an ample line bundle $E \to A$.

The second inequality of Lemma 3.1 implies

$$\begin{aligned} T_{f^{(1)},H^{(1)}}(r) &= \int_0^{2\pi} \log(1 + ||f'(re^{i\theta})||^2)\frac{d\theta}{2\pi} \\ &= m_{f^{(1)},\infty}(r) \le O(\log^+(r\,T_{f,E}(r)))/\!/\ . \end{aligned}$$

More generally, the estimate

$$T_{f^{(i)},H^{(i)}}(r) \le O(\log^+(r\,T_{f,E}(r)))/\!/ \tag{9}$$

holds for each $i \ge 1$.

The first inequality of Lemma 3.1 implies that

$$m_{f^{(i)},Z^{(i)}}(r) \le m_{f,Z}(r) + O(\log^+(r\,T_{f,E}(r)))/\!/ \tag{10}$$

holds for any $i \geq 1$. The estimate (10) does not give any information on the asymptotic behavior of $m_{f^{(1)},D^{(1)}}(r)$, or more generally, of $m_{f^{(i)},D^{(i)}}(r)$. However, if we take i sufficiently large, then we can use (10) (combined with (9)) to psove an estimate of the form

$$m_{f^{(i)},D^{(i)}}(r) \leq \varepsilon \, T_{f,E}(r) /\!/$$

(ε is any positive number and $E \to A$ is any fixed ample line bundle). We now prove this.

It is easy to see that $\dim A^{(k)} = 2^k n$ and $\dim D^{(k)} = 2^k(n-1)$. It follows that $D^{(k)}$ has codimension 2^k. There is a natural fibration $D^{(k)} \to D$. We write the fiber over $x \in D$ by $D_x^{(k)}$. Fix an arbitrary point $a \in A$. If $2^k > n$, we have

$$\begin{aligned}
(2^k - 1)n &= \text{relative dimension of } A^{(k)} \to A \\
&> 2^k(n-1) = (2^k-1)(n-1) + (n-1) \qquad [\text{since } 2^k > n] \\
&= \dim(D^{(k)})_x + \dim D \\
&= \dim\Bigg(\bigcup_{x \in D} \underbrace{D_x^{(k)}}_{\text{translated to a fixed point } a} \Bigg) .
\end{aligned}$$

This implies that if $2^k > n$ there exists a proper subvariety $\widetilde{D}^{(k)}$ in the fiber of $A^{(k)} \to A$ over a such that $D^{(k)} \subset (\widetilde{D}^{(k)} \times A)$, i.e., $D^{(k)}$ is contained in a "horizontally flat" proper subvariety. Set

$$\widetilde{D}^{(k)} = \bigcup_{x \in D} \underbrace{D_x^{(k)}}_{\text{translated to a fixed point } a} .$$

We suppose that $f^{(k)}(\mathbb{C}) \not\subset (\widetilde{D}^{(k)} \times A)$. Let E_k be a divisor in the fiber of $A^{(k)} \to A$ over a chosen so that $\widetilde{D}^{(k)} \subset E_k$ and $f^{(k)}(\mathbb{C}) \not\subset (E_k \times A)$. We then have the following estimate:

$$\begin{aligned}
m_{f^{(k)},D^{(k)}}(r) &\leq m_{f^{(k)},\widetilde{D}^{(k)} \times A}(r) \leq m_{f^{(k)},E_k \times A}(r) \\
&\leq \exists\, C \sum_{i=1}^{k} T_{f^{(i)},H^{(i)}}(r) /\!/ \\
&\leq O(\log^+(r\, T_{f,E}(r))) /\!/ \qquad [\text{from } (9)] ,
\end{aligned}$$

where the constant $C > 0$ depends only on the degree of E_k in any sense in the fiber of $A^{(k)} \to A$, and the degree of E_k with the above property depends on the degree of $\widetilde{D}^{(k)}$. For the above argument to make sense, it is sufficient that the non-inclusion $f^{(k)}(\mathbb{C}) \not\subset (\widetilde{D}^{(k)} \times A)$ holds. In [K2] I encountered the same situation. The argument given in [K2, (2.14), p.147] was not correct in the sense that I overlooked this problem (i.e., I ignored constant terms involved in the Lemma on logarithmic derivative). I would like to take this opportunity to remedy this error.

Let $\widetilde{D}^{(k)}$ be defined as above and we consider a horizontal cycle $\widetilde{D}^{(k)} \times A$ in $A^{(k)} (\to A)$. We assume that the inclusion $f^{(k)}(\mathbb{C}) \subset (\widetilde{D}^{(k)} \times A)$ holds for all k. We would like to show that if this happens then the holomorphic curve $f : \mathbb{C} \to A$ is not Zariski dense.

We solve this problem by estimating how strongly the jet lift $f^{(k)} : \mathbb{C} \to A^{(k)}$ can approximates $D^{(k)}$ for k large (roughly of order $\log_2 n$). For this purpose, we try to "deform" the horizontal cycle $\widetilde{D}^{(k)} \times A$ containing $\widetilde{D}^{(k)} \times D$, to non-horizontal ones containing $\widetilde{D}^{(k)} \times D$ but not containing the image $f^{(k)}(\mathbb{C})$. We will do this in the category of $\mathbb{Q}$-cycles.

We first consider a small non-horizontal "deformation" of the horizontal divisor $E_k \times A$ (as a $\mathbb{Q}$-divisor). We perform such small "deformation" by introducing a vertical divisor, which means a pull back F of some ample divisor of A under the projection $A^{(k)} \to A$, and consider a $\mathbb{Q}$-divisor $(E_k \times A) + \varepsilon\, F$ on $A^{(k)}$, where ε is an arbitrary positive rational number (which should be chosen to be small)[6]. Let d be a large positive integer such that $L_{k,d,\varepsilon} := \mathcal{O}_{A^{(k)}}(d\{(E_k \times A) + \varepsilon\, F\})$ is a very ample divisor on $A^{(k)}$. Then there exists a positive integer d_0 such that d_0/d is equivalent to a uniform multiple of ε, and a holomorphic section $s_{k,d,\varepsilon}$ of $L_{k,d,\varepsilon}$ such that the ideal sheaf of the subscheme $(d - d_0)(\widetilde{D}^{(k)} \times D)$ divides that of $(s_{k,d,\varepsilon})$ (over D). Thus we get a small "deformation" (over $\mathbb{Q}$) $\frac{1}{d-d_0}(s_{k,d,\varepsilon})$ of $E_k \times A$. Choosing an appropriate number of such E_k's and taking the intersection of small "deformation" (over $\mathbb{Q}$) of these $E_k \times A$'s, we get a small non-horizontal "deformation" of $\widetilde{D}^{(k)} \times A$ in the category of $\mathbb{Q}$-cycles. Moreover, given ample divisor D (in A), $k \in \mathbb{Z}_{>0}$ and any ample line bundle E, we have $m(D, k) \in \mathbb{Z}_{>0}$ such that for any E_k (in the fiber of $A^{(k)} \to A$) containing $\widetilde{D}^{(k)}$ and of "degree $\geq m(D, k)$" w.r.to the ample line bundle E, we may assume that $\widetilde{D}^{(k)} \times A$ and its small "deformation" (over $\mathbb{Q}$) are disjoint in the fiber of $A^{(k)} \to A$ over some point of A. We now would like to expect the non-inclusion $f^{(k)} \not\subset \mathrm{Supp}((s_{k,d,\varepsilon})_0)$.

However, we cannot expect this non-inclusion because of the following reason. Namely, the section $s_{k,d,\varepsilon} \in H^0(A^{(k)}, \mathcal{O}(L_{k,d,\varepsilon}))$ is special in the sense that $(d - d_0)(\widetilde{D}^{(k)} \times D)$ divides $(s_{k,d,\varepsilon})$ over D. So, it might happen that the divisor $(s_{k,d,\varepsilon})$ inherits this special property, if ε (the "strength" of the perturbation) is very small (and we want arbitrarily small ε). In other words, we have two requirements, i.e.,

(a) to make ε in the perturbation arbitrarily small, and

(b) to find such ε in the perturbation so that the perturbed linear system $|d\{(E_k \times A) + \varepsilon\, F\}|$ contains a linear subsystem having exactly $(d - d_0)(\widetilde{D}^{(k)} \times D)$ as the base scheme.

However, there may not exist such $\varepsilon > 0$ which fulfills both conditions (a) and (b).

Therefore, the worst such case is that there exists a proper subvariety Y of A such that any divisor $(s_{k,d,\varepsilon})$ having the above divisibility property necessarily contains a subscheme of the form $Z_k \times Y$ (Z_k being a subvariety of $\widetilde{D}^{(k)}$) and the inclusion $f^{(k)}(\mathbb{C}) \subset Z_k \times Y$ holds. We cannot avoid this possibility in the small "deformation" of $E_k \times A$ (see the discussion after Theorem 3.1 below). Therefore, instead of looking for a "good" $E_k \times A$, we now put the assumption that the assumption the holomorphic curve $f : \mathbb{C} \to A$ is **Zariski-dense (algebraically non-degenerate)**.

Under this assumption, we have the estimate

$$m_{f^{(k)},(d-d_0)\, D^{(k)}}(r) \leq m_{f^{(k)},(s_{k,d,\varepsilon})}(r)\ .$$

[6] To make the error term $\varepsilon\, T_{f,E}(r)$ in (1) smaller, we need to choose $\varepsilon > 0$ accordingly smaller

Dividing the both sides by $d - d_0$ we have

$$m_{f^{(k)},D^{(k)}}(r) \leq \frac{d}{d-d_0} T_{f^{(k)},(E_k \times A)+\varepsilon F}(r) \ .$$

If we choose ε sufficiently small and d sufficiently large, we can make d_0/d arbitrarily small. Combining this with (9), we finally have the estimate of type

$$m_{f^{(k)},D^{(k)}}(r) \leq \varepsilon\, T_{f,E}(r) /\!/ \ ,$$

where E is any fixed ample line bundle on A and ε is any positive number. We have thus proved the main result in Yamanoi's paper [Y2] (the Second Main Theorem for Zariski-dense holomorphic curves into Abelian varieties) by studying the truncated counting functions in the framework of the Radon transformation.

Theorem 3.1 (Second Main Theorem with truncation level 1). *Let A be an Abelian variety, D any reduced divisor and ε any small positive number. Let $f : \mathbb{C} \to A$ be a Zariski-dense holomorphic curve. Then we have*

$$m_{f,D}(r) + N^1_{f,D}(r) \leq \varepsilon\, T_{f,E}(r) /\!/ \ .$$

The assumption $f : \mathbb{C} \to A$ being Zariski dense cannot be removed. Indeed, we choose an Abelian variety A which admits a non-trivial proper Abelian subvariety $i : B \hookrightarrow A$ and a reduced divisor D of A having the property that $i^*D = m\, D_B$ for $\mathbb{Z} \ni m \geq 2$ and D_B is a reduced divisor on B. Let $f : \mathbb{C} \to A$ be a holomorphic curve such that $f(\mathbb{C}) \subset B$. This choice of A and f violates the assumption of Theorem 3.1. On the other hand, since all intersections of f and D are with multiplicity ≥ 2, we have $N^1_{f,D}(r) \geq \frac{1}{2} N_{f,D}(r)$. This implies

$$\begin{aligned} m_{f,D}(r) + N^1_{f,D}(r) &\geq \frac{1}{2} m_{f,D}(r) + \frac{1}{2}(m_{f,D}(r) + N_{f,D}(r)) \\ &= \frac{1}{2} m_{f,D}(r) + \frac{1}{2} T_{f,D}(r) + O(1) \\ &\geq \frac{1}{2} T_{f,D}(r) \ . \end{aligned}$$

This shows that the Second Main Theorem with the level 1 residual counting function does not hold for this choice of A and f. Therefore the assumption f being Zariski dense cannot be removed.

To sum up, we have proved the following: if $f : \mathbb{C} \to A$ is Zariski-dense, then we conclude that the intersection points of f and D with high multiplicity are "rare" in the sense that Theorem 3.1 holds.

References

[**A**] L. Ahlfors, *The theory of meromorphic curves*, Acta Soc. Sci. Finn. N.S. A. **III** (1941), 1-31.

[**B**] M. Brunella, *Courbe entière et feulletages holomorphes*, Enseign. Math (2) **45** (1999), 195-216.

[**F**] H. Fujimoto, *Value Distribution Theory of the Gauss Map of Minimal Surfaces in $\mathbb{R}^n$*, Vieweg, 1993.

[**C**] H. Cartan, *Sur les zéros des combinaisons linéaires de p fonctions holomorphes donnés*, Mathematica **7** (1933), 5-31.

[**G-G**] M. Green and P. A. Griffiths, *Two applications of algebraic geometry to entire holomorphic mappings*, Chern Symposium, Springer-Verlag, 1980, pp. 41-74.

[**K1**] R. Kobayashi, *Nevanlinna Theory from the View Point of Lemma on Logarithmic Derivative*, (in Japanese) to be published in "Surveys in Geometry, Sp. Ed." (2005), Baifukan.

[**K2**] R. Kobayashi, *Holomorphic curves in Abelian varieties: The second main theorem and applications*, Japanese J. Math. (2000), 129-152.

[**K3**] R. Kobayashi, *Toward Nevanlinna theory as a geometric model of Diophantine approximation*, AMS Sugaku Expositions **16** (2003), 39-79.

[**KS**] S. Kobayashi, *Hyperbolic Complex Spaces*, Springer-Verlag, 1998.

[**L1**] S. Lang, *Number Theory III*, Encyclop. Math. Sc., Vol. 60, Springer-Verlag, 1991.

[**L2**] S. Lang, *Introduction to Complex Hyperbolic Spaces*, Springer-Verlag, 1987.

[**M**] M. McQuillan, *A Dynamical counterpart to Faltings' "Diophantine approximation on Abelian varieties"*, IHES preprint (1996).

[**N**] R. Nevanlinna, *Analytic Functions*, Springer-Verlag, 1970.

[**NoWY**] J. Noguchi, J. Winkelmann and K. Yamanoi, *The second main theorem for holomorphic curves in semi-Abelian varieties*, preprint (1999).

[**NoO**] J. Noguchi and T. Ochiai, *Geometric Function Theory in Several variables*, Transl. Math. Monog. 80, AMS, 1990.

[**O**] T. Ochiai, *On holomorphic curves in algebraic varieties with ample irregularity*, Invent. Math. **44** (1977), 83-96.

[**R**] M. Ru, *A defect relation for holomorphic curves intersecting hypersurfaces*, preprint (2004).

[**S**] J. Silverman, *Arithmetic distance functions and height functions in Diophantine geometry*, Math. Ann. **279** (1987), 193-216.

[**SY**] Y.T.Siu and S.K.Yeung, *Defects for ample divisors of Abelian varieties, Schwarz lemma and hyperbolic hypersurfaces of low degrees*, Amer. J. Math. **119** (1997), 1139-1172.

[**V1**] P. Vojta, *Diophantine Approximation and Value Distribution Theory*, LNM 1239, Springer Verlag, 1987.

[**V2**] P. Vojta, *A refinement of Schmidt subspace theorem*, Amer. J. Math. **111** (1989), 489-518.

[**V3**] P. Vojta, *On Cartan's theorem and Cartan's conjecture*, Amer. J. Math. **119** (1997), 1-17.

[**W**] J. Weyl, *Meromorphic curves*, Ann. Math. **39** (1938), 516-538.

[**Y1**] K. Yamanoi, *Algebro-geometric version of Nevanlinna's lemma on logarithmic derivative and applications*, to appear in Nagoya Math. Journ. **174** (2004), 23-63.

[**Y2**] K. Yamanoi, *Holomorphic curves in Abelian varieties and intersections with higher codimensional subvarieties*, Forum Math. **16** (2004), 749-788.